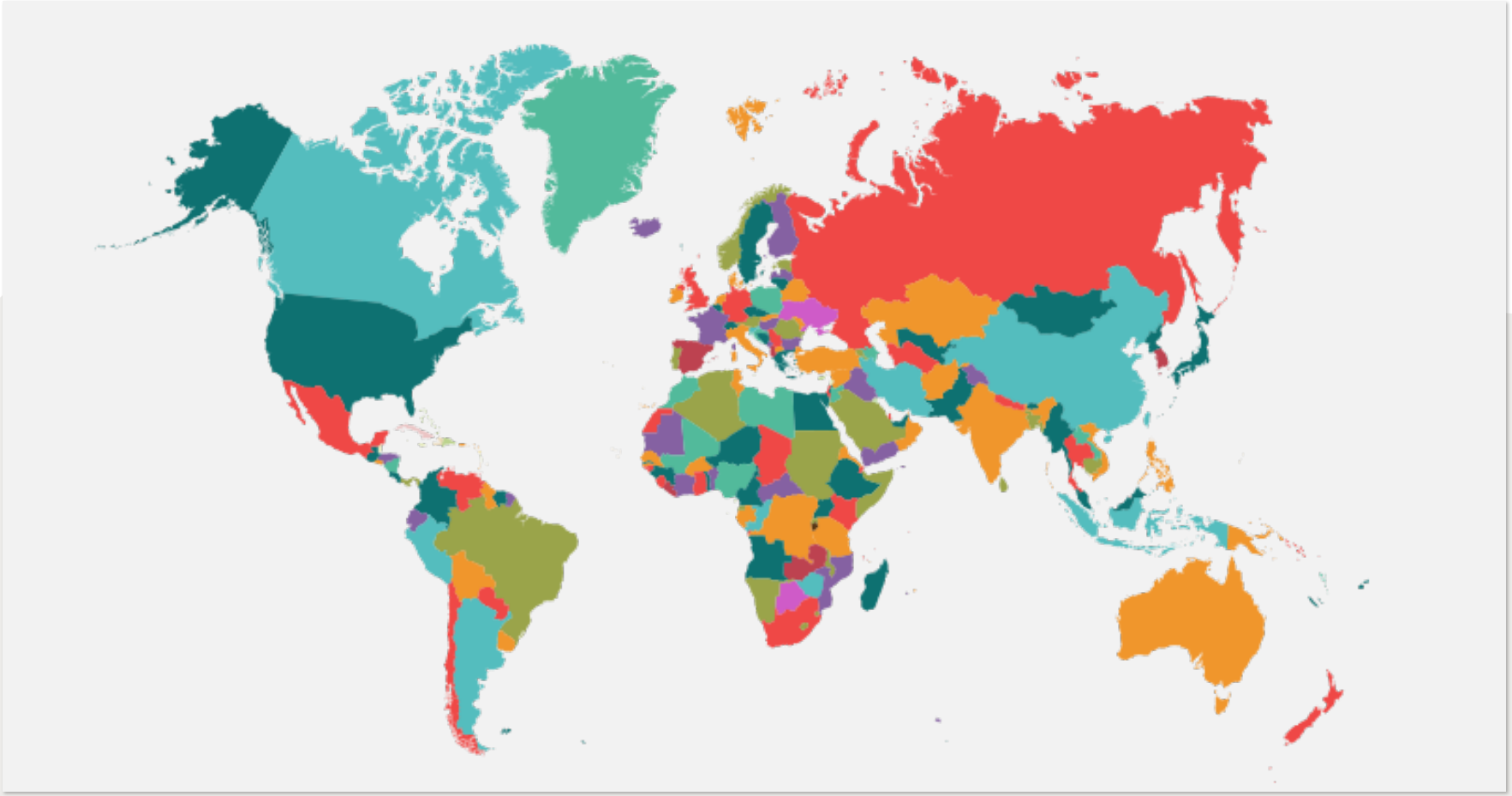


Country Intelligence Web Mining

The Dashbordians Team

Dashboarding

Andrew Piasecki

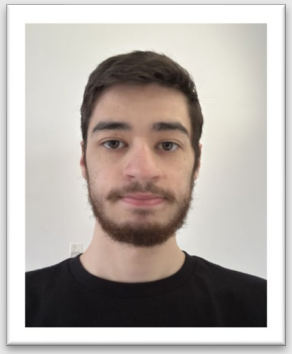
Analysis & Operations

Henry Gillis

Kathryn James

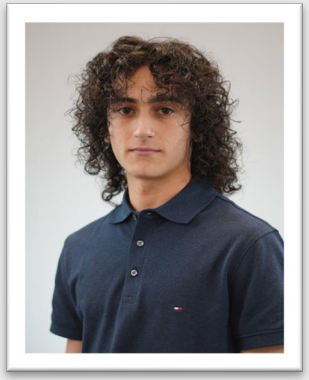
Brandon Connell

Scripting

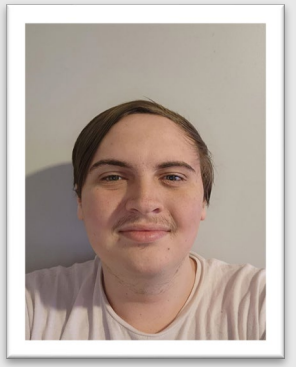


Michael Zaino

Ethan McWhirter



Gustavo Salama

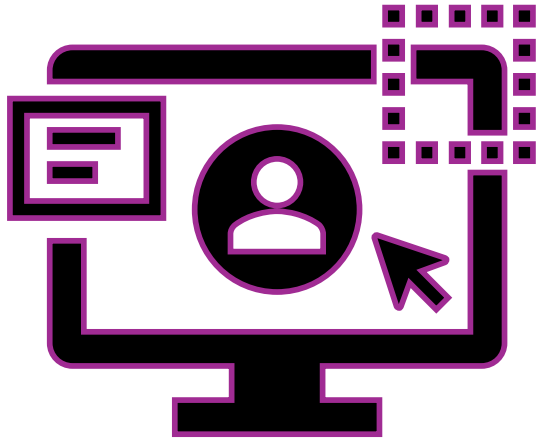


Ethan Davis

Problem Statement

Automate the process of gathering and analyzing country intelligence from public web sources.

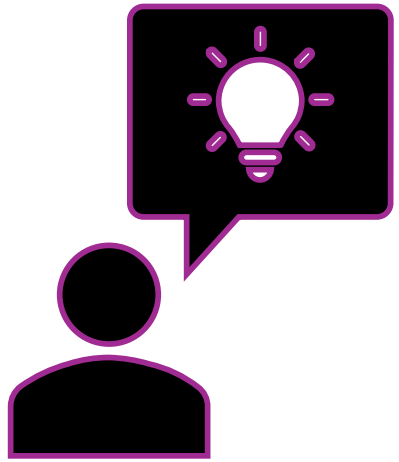
Customers' Main Goal



Automate the process of gathering and analyzing country intelligence from public web sources.

- Our solution: Design a system that can efficiently and effectively scrape data and compile it into user-friendly dashboards

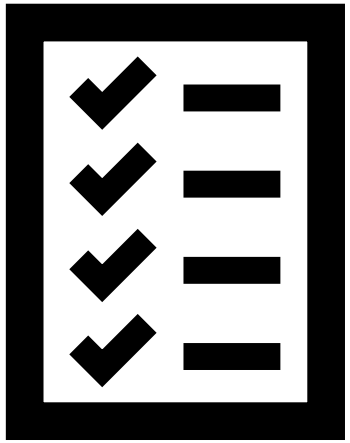
Customer Bonus Objectives



Find or create new ways to gain insights that could be used by our sponsor

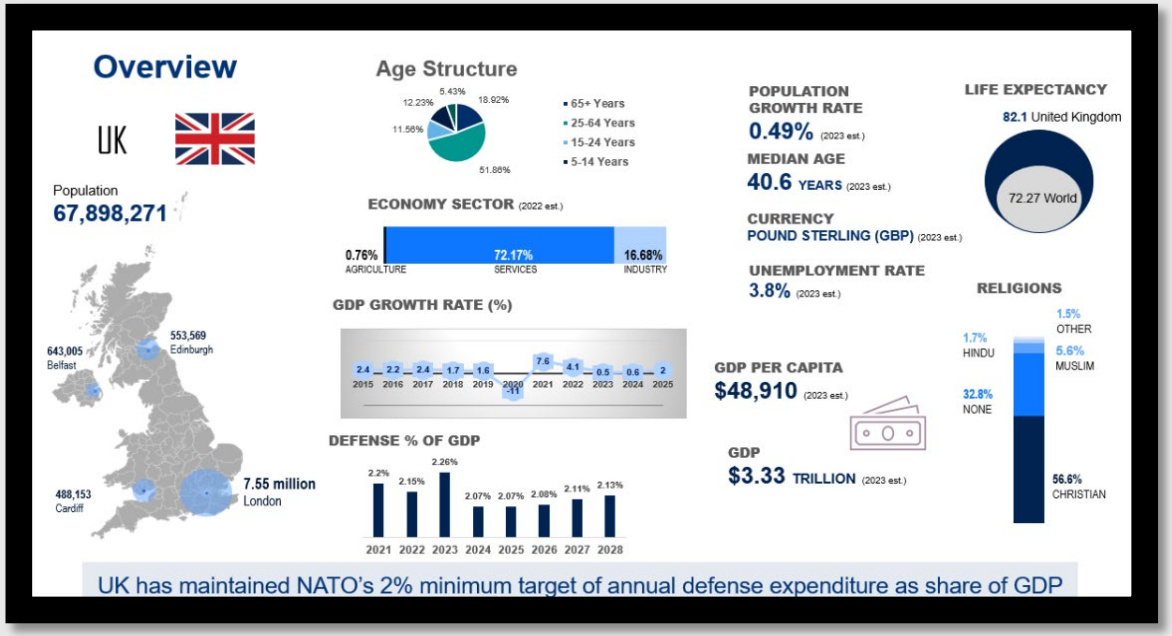
- Our solution: A web application that integrates with gemma3:1b for AI-powered responses and Sentiment Analysis

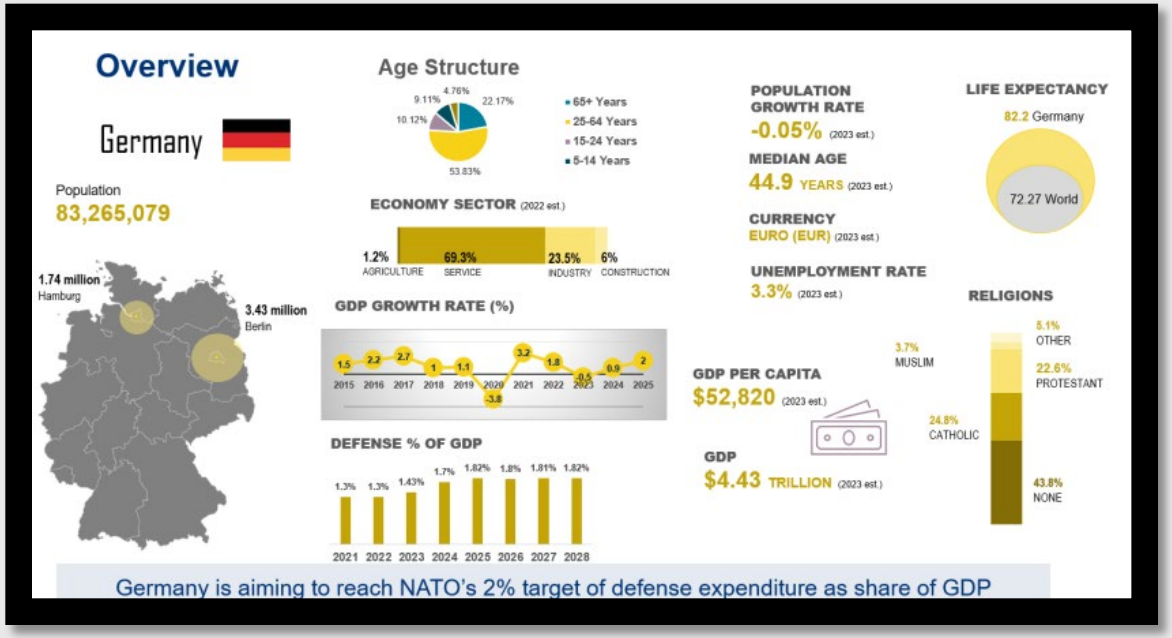
Deliverables



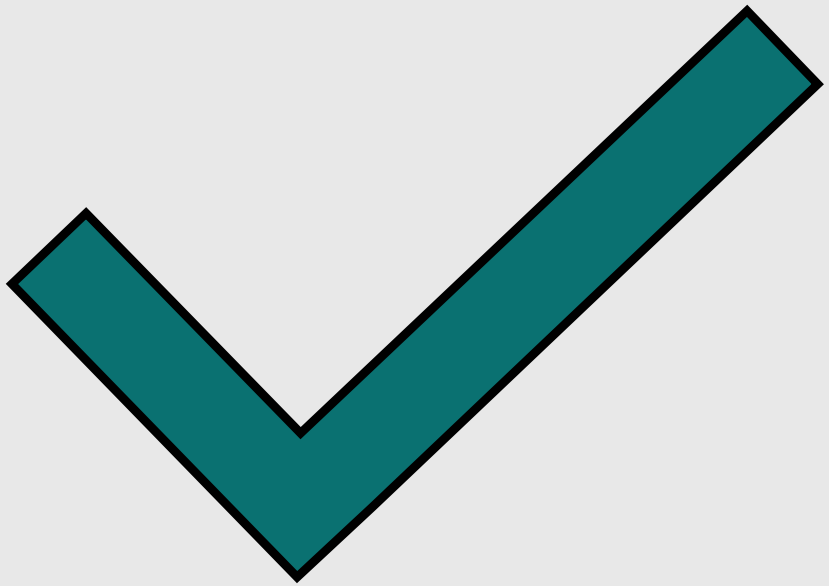
Desirables

- Compile data into accessible information visually and numerically
  - GDP, population, defense data
  - Used for presentations
  - General evaluation of collected data for other decision-making processes
- Why Tableau?
  - They already utilize it for other tasks
  - Better than Power BI for utilizing 20-30+ sheets
- Must use reputable sources
  - Periodic web scraping
    - Updated current information





Dashboards



Part 1 Completion

Dashboard creation, the essential component of the project to satisfy customers' needs, was completed on 01/19/2025

M.A.R.T.I.N

Multi-source Article Ranking & Textual Insights Navigation

MARTIN was created to streamline the digestion of current events, specifically tailored to key domains like defense and conflicts.

Overview:

- A web app that integrates with gemma3:1b using Ollama for AI-powered responses.

How it works:

1. User submits a query
2. Relevant articles are retrieved
3. gemma3:1b processes the data & generates a response
4. Response is displayed on the website

Key Enhancements Coming:

◆ Improve UI/UX ◆ Enhance response accuracy ◆ Make it online

M.A.R.T.I.N

MARTIN - Advanced Article Analysis

Chat | Rankings

MARTIN Assistant:

Ask MARTIN about current events, military news, or any other topic. The assistant uses advanced ML models to analyze content for combat, monetary, and political relevance.

Hello there! I'm here to assist you. Do you have any specific questions or topics you'd like me to delve into?

yes, please tell me about the article with the highest combat rank

You can read more about this here: https://www.twz.com/air/boeing-wins-air-forces-next-generation-air-dominance-fighter-contract

give a summary about https://www.twz.com/air/boeing-wins-air-forces-next-generation-air-dominance-fighter-contract

Okay, here's a summary of the Boeing contract, based on the provided context: Boeing has won a significant contract to develop and produce the F-47, a new crewed sixth-generation stealth combat jet, as the centerpiece of the U.S. Air Force's Next Generation Air Dominance (NGAD) program. This contract is generating considerable buzz due to its potential to dramatically shift air combat capabilities. The contract is viewed as a major advancement for the U.S. military, with the Air Force hoping to begin fielding the F-47 in the coming decade. It is also noted that the F-47 has received mixed sentiment, with some experts expressing concerns about its origins and the potential implications of the deal.

Type your message here... Send

MARTIN - Advanced Article Analysis

Chat | Rankings

Understanding the Rankings:

Each article is analyzed using advanced machine learning techniques to score relevance across different aspects:

- **Combat Rank (0-100):** Measures relevance to actual ongoing combat situations, distinguishing between real conflicts and military exercises/tests.
- **Monetary Rank (0-100):** Indicates relevance to financial and economic topics.
- **Political Rank (0-100):** Shows relevance to political and governmental matters.
- **Sentiment Score (-100 to +100):** Indicates the article's emotional tone, from very negative to very positive.

Title	Combat Rank	Monetary Rank	Political Rank	Sentiment Score	Link
Bunker Talk: Let's Talk About All The Things We Did And Didn't Cover This Week Bunker Talk: Let's Talk About All The Things We Did And Didn't Cover This Week	4.2	0.1	10.0	98.5	Read Article
Israel-Turkey Tensions Mounting Over Air Bases In Syria Israel-Turkey Tensions Mounting Over Air Bases In Syria	59.3	2.7	30.1	-94.0	Read Article
Coyote, Roadrunner Loitering Drone Interceptors To Arm U.S. Navy Destroyers Coyote, Roadrunner Loitering Drone Interceptors To Arm U.S. Navy Destroyers	**73.9**	28.4	21.5	-94.3	Read Article
Signal For Secure Comms: Convenience Over Security Without The Record-Keeping Signal For Secure Comms: Convenience Over Security Without The Record-Keeping	17.7	0.2	**97.6**	-95.4	Read Article
What Cat Name Should Northrop Grumman Give Its Fighter If It Wins The Navy's F/A-XX Competition? What Cat Name Should Northrop Grumman Give Its Fighter If It Wins The Navy's F/A-XX Competition?	20.3	68.8	**74.5**	-95.4	Read Article
Turkish F-16s Are Using Tablets To Control Locally Made Weapons Turkish F-16s Are Using Tablets To Control Locally Made Weapons	2.5	0.8	0.7	-93.5	Read Article
Reflections Of The F-47: Looking Back At The X-36 Reflections Of The F-47: Looking Back At The X-36	3.6	29.4	11.1	-95.0	Read Article
ME-11B Official Designation Of Army's New Intelligence Gathering Business Jets ME-11B Official Designation Of Army's New Intelligence Gathering Business Jets	13.8	39.4	56.3	-90.9	Read Article
First North Korean Airborne Early Warning Jet Flies, Kim Shows Off Interior First North Korean Airborne Early Warning Jet Flies, Kim Shows Off Interior	2.7	28.5	**77.0**	92.6	Read Article

☑ Customer Needs Met

Part 1

- Various dashboards were created to satisfy the component of *"…gathering and analyzing country intelligence from public web sources."*
  - Country overview
  - GFP
  - Physical map

→

Part 2

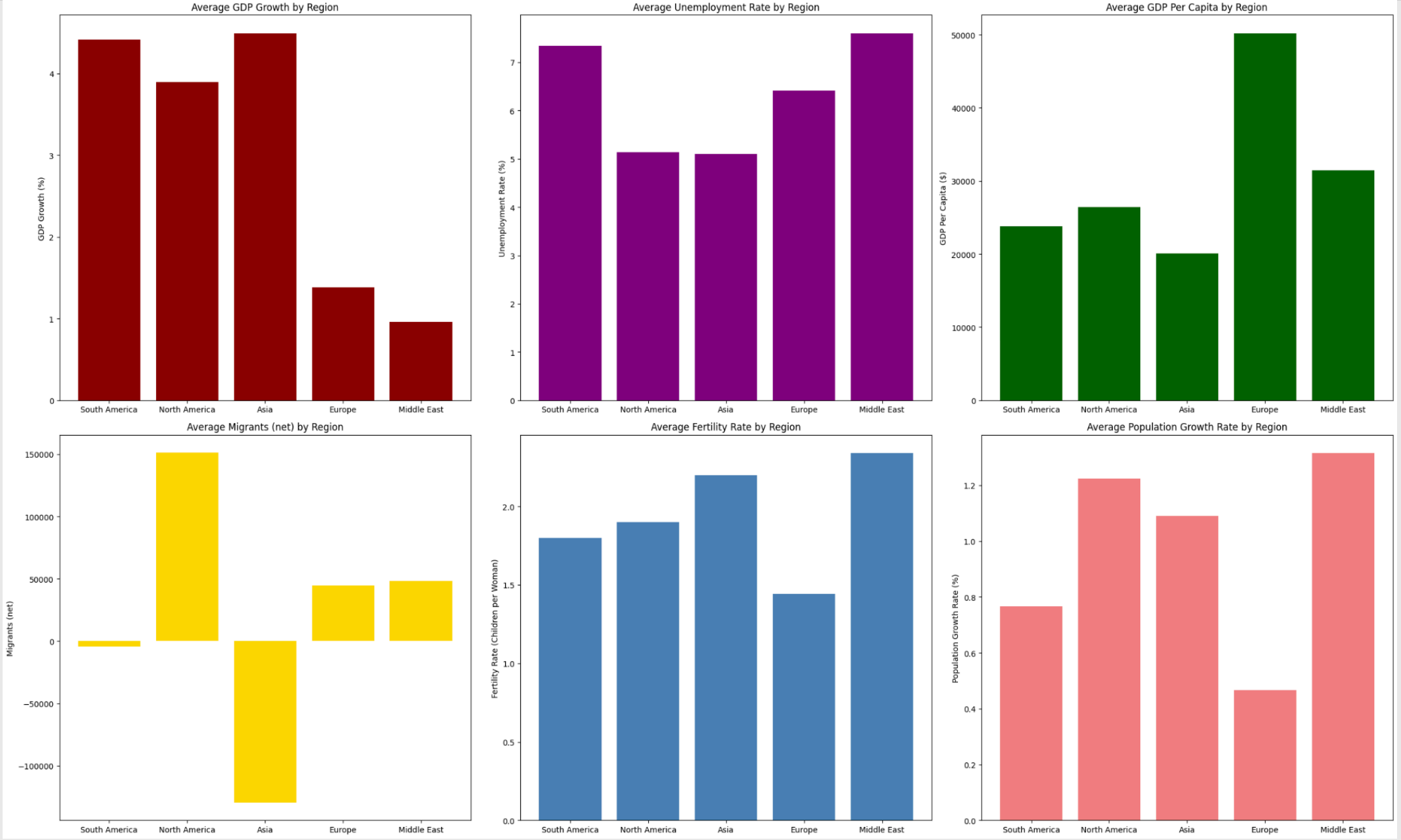
- Ability to ingest breaking news articles
- Chatbot trained on said articles
- Rating the relevance of articles to certain topics of interest

Exploratory Analysis

Exploratory Analysis of Country Data

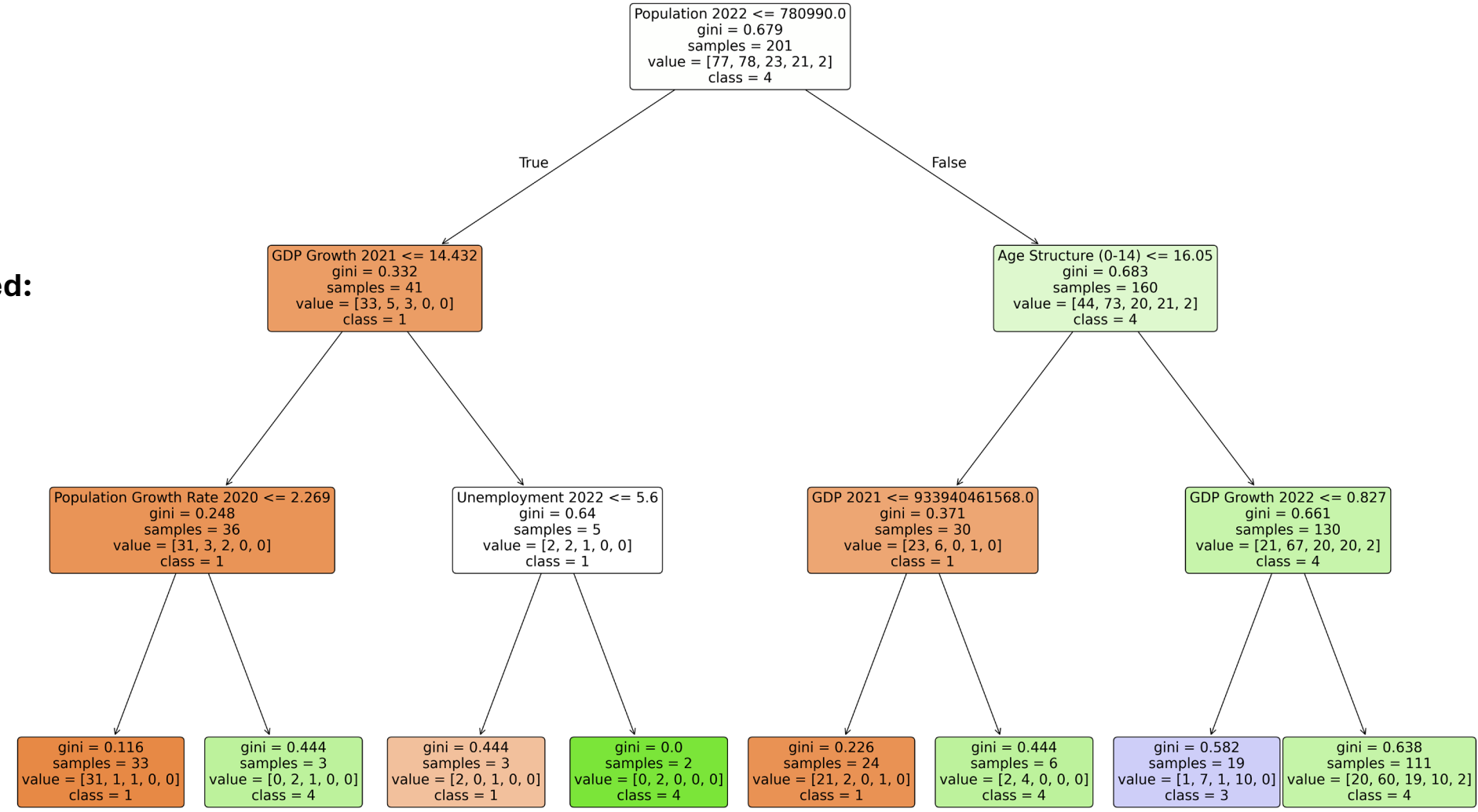
- **Regional Breakdown:**
  - South America, North America, Europe, Middle East, Asia
- **Comparative Insights:**
  - Economic strength, population trends, inflation patterns
- **Correlation Analysis:**
  - GDP per capita, population, inflation
- **Predictive Analysis:**
  - Can we predict U.S. Department of State advisory levels?

Comparative Analysis of Country Data



Predictive Analysis: U.S. Travel Advisories

- **Random Forest model:**
  - 47% accuracy
- **Decision Tree model:**
  - Easier to interpret
  - Worse performance
- **Other Potential models explored:**
  - Predicting the time of a change
  - Using Wikipedia edit counts as predictors


Population 2022 <= 780990.0
gini = 0.679
samples = 201
value = [77, 78, 23, 21, 2]
class = 4
True
False
GDP Growth 2021 <= 14.432
gini = 0.332
samples = 41
value = [33, 5, 3, 0, 0]
class = 1
Age Structure (0-14) <= 16.05
gini = 0.683
samples = 160
value = [44, 73, 20, 21, 2]
class = 4
Population Growth Rate 2020 <= 2.269
gini = 0.248
samples = 36
value = [31, 3, 2, 0, 0]
class = 1
Unemployment 2022 <= 5.6
gini = 0.64
samples = 5
value = [2, 2, 1, 0, 0]
class = 1
GDP 2021 <= 933940461568.0
gini = 0.371
samples = 30
value = [23, 6, 0, 1, 0]
class = 1
GDP Growth 2022 <= 0.827
gini = 0.661
samples = 130
value = [21, 67, 20, 20, 2]
class = 4
gini = 0.116
samples = 33
value = [31, 1, 1, 0, 0]
class = 1
gini = 0.444
samples = 3
value = [0, 2, 1, 0, 0]
class = 4
gini = 0.444
samples = 3
value = [2, 0, 1, 0, 0]
class = 1
gini = 0.0
samples = 2
value = [0, 2, 0, 0, 0]
class = 4
gini = 0.226
samples = 24
value = [21, 2, 0, 1, 0]
class = 1
gini = 0.444
samples = 6
value = [2, 4, 0, 0, 0]
class = 4
gini = 0.582
samples = 19
value = [1, 7, 1, 10, 0]
class = 3
gini = 0.638
samples = 111
value = [20, 60, 19, 10, 2]
class = 4

Conclusion and Future Recommendations

- Customer needs were successfully met
  - Dashboard creation
  - MARTIN developments
  - 30% reduction of hours spent per study (40 -> 28 hours)

- Recommendations
  - Further analysis of country data
  - Further improvements of MARTIN
  - Additional information gathering tools

Questions?